(19) World Intellectual Property Organization International Bureau



(43) International Publication Date 25 October 2001 (25.10.2001)

(10) International Publication Number WO 01/79449 A2

(51)	International Patent Classification?	C12N	Jose, CA 95117 (US). DRMANAC, Radoje, T. [YU/US]; 850 East Greenwich Place, Palo Alto, CA 94303 (US).
(21)	International Application Number:	PCT/US01/08656	
(22)	International Filing Date: 16 Apr	il 2001 (16.04.2001)	(74) Agent: ELRIFI, Ivor, R.; Mintz, Levin, Cohn, Ferris, Glovsksy and Popeo, P.C., One Financial Center, Boston, MA 02111 (US).
(25)	Filing Language:	English	
(26)	Publication Language:	English	(81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM,

18 April 2000 (18.04.2000)

26 January 2001 (26.01.2001)

(63) Related by continuation (CON) or continuation-in-part (CIP) to earlier applications:

US 09/552,929 (CIP) Filed on 18 April 2000 (18.04.2000) US 09/770,160 (CIP) Filed on 26 January 2001 (26.01.2001)

(71) Applicant (for all designated States except US): HYSEQ, INC. [US/US]; 670 Almanor Avenue, Sunnyvale, CA 94086 (US).

(72) Inventors; and

(30) Priority Data:

09/552,929

09/770,160

(75) Inventors/Applicants (for US only): TANG, Y., TOM [US/US]; 4230 Ranwick Court, San Jose, CA 95118 (US). LIU, Chenghua [CN/US]; 1125 Ranchero Way #14, San

- nd Popeo, P.C., One Financial Center, Boston,
- States (national): AE, AG, AL, AM, AT, AU, B, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, L, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.
- (84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:

without international search report and to be republished upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES

(57) Abstract: The present invention provides novel nucleic acids, novel polypeptide sequences encoded by these nucleic acids and uses thereof.

No: of N		SEQ ID	SEQ ID	IM-	CEO ID NO.	Nucleatide	Nucleotide	IA mino gold convenes (V-U-U +-C)
d d d d d d d d d d						1		Amino acid sequence (X=Unknown; *=Stop
				٦	05.770,100			nociconae inscritorij
			,	1	}			j
TSPSPGTRAASQPPNSSKAGRAPW DRNNPLRNPLSSNLVRNPLLAKGPF KLRAPFSQQPHSKMKPAGSVSDM LDAFDLDRMKQEI*KEVVRELHK GERKEIDMARQEA*SGISMKNHAG GERKEIDMARQEA*SGISMKNHAG GERKEIDMARQEA*SGISMKNHAG GERGVARHDRPRGTLREYKVVGRK LTYFK/CHTPPL/YRMRIFAPPSMSSI SPRFWYFVSQLKKDEESLQWRFSY CAQVFEKSPSLRVKNHGFUNKNYDS RSGYTHNMYREYVRDLDHPQAPVH SCLTRDNOVAPAPAA/HEAHFHFRE ERLEEILAGQQDCRRPGCSKQFPRS RFKFPAAPPOSILRRQQDCRRPGCSKQFPRS RFKFPAAPPOSILRRQQDCRRPGCSKQFPS RFKFPAAPPOSILRRQQDCRRPGCSKQFPS RFKFPAAPPOSILRRQQDCRRPGCSKQFPS RFKFPAAPPOSILRRQQDCRRPGCSKQFS RFKFPAAPPOSILRRQQDCRRPGCSKQFT RFKFPAAPPOSIT RFKFPAAP		<u>_</u> _	.					
TSPSPGTRAASQPPNSSKAGRAPW DRNNPLRNPLSSNLVRNPLLAKGPF KLRAPFSQQPHSKMKPAGSVSDM LDAFDLDRMKQEI*KEVVRELHK GERKEIDMARQEA*SGISMKNHAG GERKEIDMARQEA*SGISMKNHAG GERKEIDMARQEA*SGISMKNHAG GERGVARHDRPRGTLREYKVVGRK LTYFK/CHTPPL/YRMRIFAPPSMSSI SPRFWYFVSQLKKDEESLQWRFSY CAQVFEKSPSLRVKNHGFUNKNYDS RSGYTHNMYREYVRDLDHPQAPVH SCLTRDNOVAPAPAA/HEAHFHFRE ERLEEILAGQQDCRRPGCSKQFPRS RFKFPAAPPOSILRRQQDCRRPGCSKQFPRS RFKFPAAPPOSILRRQQDCRRPGCSKQFPS RFKFPAAPPOSILRRQQDCRRPGCSKQFPS RFKFPAAPPOSILRRQQDCRRPGCSKQFPS RFKFPAAPPOSILRRQQDCRRPGCSKQFS RFKFPAAPPOSILRRQQDCRRPGCSKQFT RFKFPAAPPOSIT RFKFPAAP				1	1			RKAASOSDKPAEKKEDESOMEDPS
DRNNPLRIPLESSNLVRPILLAKOP KLRAPFSQQPHSRMKPAGSIVSDM/ LDAFDLDLRMKQEI*KEVVRELHK GERKEIIDVAIRQEA*SGISRKKNLGI RAHPFITSFICSQPRILM GERKEIIDVAIRQEA*SGISRKKNLGI RAHPFITSFICSQPRILM GERGVARHDRPGTTLREYKVVGRE LPTPK\CHTPFL\YRIMRIFAP*SIMSI SPRFWYFVSQLKKDEESLQWFFSY CAQVFEKSPLRVKINGGIWLRIYDS RSGITHNMYREVRBLDHPQAPVH SCLTRDNGVAPAPAA/HEAHFHFRE ERLEEINAGQOCRENPGSKQFPRIS RRFKPAAPPGSLRRQDKPRFITTKRE ERLEEINAGQOCRENPGSKQFPRIS RRFKPAAPPGSLRRQDKPRFITTKRE ERLEEINAGQOCRENPGSKQFPRIS RRFKPAAPPGSLRRQDKPRFITTKRE ERLEEINAGQOCRENPGSKGFPRIS RRFKPAAPPGSLRRQDKPRFITTKRE ERLEEINAGQOCRENPGSKGFPRIS RRFKPAAPPGSLRRQDKPRFITTKRE ERLEEINAGQOCRENPGSKGFPRIS RRFKPAAPPGSLRRQDKFRFITTKRE ERLEEINAGQOCRENPGSKGFPRIS RRFKPAAPPGSLRRQDKSVPRITTACIDVFGWID HESIMNVNVKGVAWGGDKARW GNEDQKEGQEGKRSLSIEHLLCSGP SDFAAHFVQLGELKAAFSIDEKTRE EQ SPANTYVISPLSENFSARYRNHSND LTCVHTELQNINTKLTVLEGDILDEP FLKRACQDVSVNIHTACIDVFGWID HRESIMNVNVKGRVAWGGDKARW GNEDQKEGQEGKRSLSIEHLLCSGP SDFAAHFVQLGELKAAFSIDEKTRE EQ SPANTYVISPLSEDLLPSIFCFLVSD SCYFGLATMGWSCLVTGAGGLGC RIVRLLVEGGLLPSIFCFLVSD SCYFGLATMGWSCLVTGAGGLGC RIVRLLVEGGULPSIFCFLVSD SCYFGLATMGWSCLVTGAGGLGC RIVRLLVEGGULPSIT EQ A 1881 53 1338 CPLQGHPRVTLESDLLPSIFCFLVSD SCYFGLATMGWSCLVTGAGGLGC RIVRLLVEGGULPSTCH EQ A 1882 53 1338 CPLQGHPRVTLESDLLPSIFCFLVSD SCYFGLATMGWSCLVTGAGGLGC RIVRLLVEGGULPSTCH EQ A 1882 53 1338 CPLQGHPRVTLESDLLPSIFCFLVSD SCYFGLATMGWSCLVTGAGGLGC RIVRLVEGGULPSTCH EQ A 1882 53 1338 CPLQGHPRVTLESDLLPSIFCFLVSD SCYFGLATMGWSCLVTGAGGLAC RIVRLVEGGULPSTCH EQ A 1883 1 575 HSLFGTSEVINKLLVPDAMGHIFTE DIVATITSL WGKVINVEDAGGSTT GRSSLVVYPWTQRFPDSFGNLSS ASAMGKPYSKAHGKVLTTLCT MFTKHLFFRGGGGRSLATAGKW VGSWWVPPVPAQGRNLEVPRQ GRPAPEDGGGRSLFNAGKWLTTLCT MFTKHLFFRGGGGRSLATAGKW THE TIGHT MFTKHLFFRGGGGGRSLATAGKW THE TIGHT MFTKHLFFRGGGGGRSLATAGKW THE TIGHT MFTKHLFFRGGGGGRSLATAGKW THE TIGHT MFTKHLFFRGGGGRSLATAGKW THE TIGHT MFTKHLFFRGGGGRSLATAGKW THE TIGHT MFTKHLFFRGGGGGRSLATAGKW THE TIGHT MFTKHLFFRGGGGGRSLATAGKW THE TIGHT MFTKHLFFRGGGGGRSLATAGKW THE TIGHT MFTKHLFFRGGGGAGAGAGA GRANGGKW THE TIGHT MFTKHLFFRGGGGAGAGAGAG GRANGGKW T		İ		1		1	J	
		1						1
1746		ł	i			(ł	
1746 7243 A 1876 1 668 GERKUNARIDERGILREYKVVGRK CAMPPTRTSFICSQRPILM CERGVARIDERGILREYKVVGRK LPTPKUCHTPPLYRIMKIRAP'SMSSI SPRFWYFVSQLKKDEESLQWKFSY CAQVFEKSPLRVKINFGIWLRIYDS RSGITHNMYNEYNDLDHPQAPVHR SCLTRDNGVAPAPAAHEAHFIRE ERLEEIAGQQDCRRPGCSKQFPRIS RFKFPAAPPGSLRRQDKPRFITKRF ERLEEIAGQQDCRRPGCSKQFPRIS RFKFPAAPPGSLRRQDKPRFITKRF ERLEEIAGQQDCRRPGCSKQFPRIS RFKFPAAPPGSLRRQDKPRFITKRF ERLEEIAGQQDCRRPGCSKQFPRIS T749 7245 A 1878 87 260		ļ	1	1	1	1	ļ	
1746		ļ		}	j	l		
1746		ĺ			1	İ		GERKEIID\AIRQEA*SGISRKKNLGH
1746		i -		1		l		RAHPPTRTSFICSQRPRLM
LPTPKICHTPPLIVRIMRIFAP'SMSIS SPRFWYFVSQLKKDEESLQWRFSY CAQVFEKSPLRVKINFGIWLRIVDS RSGITHNMYREYYRDLDHPQAPVH SCLTRDNGVAPAPAAHEAHFHFRF ERLEEIAGQQDCRRFGCSKQFFRIS RFKPPAAPPGSLRRQDKPRTITKRR ERLEEIAGQQDCRRFGCSKQFFRIS RFKPPAAPPGSLRRQDKPRTITKRR ERLEEIAGQQDCRRFGCSKQFFRIS RFKPPAAPPGSLRRQDKPRTITKRR ERLEEIAGQQDCRRFGCSKQFFRIS RFKPPAAPPGSLRRQDKPRTITKRR T745 A 1879 1 1254 1749 7246 A 1879 1 1254 1750 7247 A 1880 160 615 PSLNTYVTSPLSENFSARYRNHSND LTCVHTELQNKTKLTVLEGDILDEP FLKRACQDVSVINITACIDDVFGVI HRESIMIVNVKGRVAWGGDKARW GNEDQKEGQEKRSLSIEHLLCSGP SDFADHYQLGELKAAIFSFIDEKTR EQ SDFADHYQLGELKAAIFSTIDEKTR EQ SDFADHYQLGELT EQ SDFA		1746	7243	Ã	1876	1	668	
SPRFWYFVSQLKKDEESLQWRFSY				'				
CAQVFEKSPLRVKINFGIWLRIVDS RSG\Thm\y\represcript{RPILDHPQAPVH} SCLTRDNOVAPAPAAHEAHFHFRE ERLEEIAGQQDCRRFGCSKQFFR18 FRFPAAPPGSLRRQDKPRFITKE KTFLKVQGFSSVCPQNKTQETPR 1748 7245 A 1878 87 260 1749 7246 A 1879 1 1254 1750 7247 A 1880 160 615 PSLNTYVTSPLSENFSARYRNHSND LTCVHTELQNKTKLTVLEGDILDEP FLKRACQDVSVINITACIDDYGGVI HRESIMNVNVKGRVAWGGDKARW GNEDQKEQEGKRSLSIEHLLCSGP SCYPGLATMGWSCLVTGAGGLLGG RIVRLLVEKELKEIRALDKAFFPEL LKESLARDRLRSIIHTACFHLMSFGV VITHREFFMNVQC*KVPSC*CEACVQ ASVPVFIYTSSIEVAGPNSYKEIIQNG HEEEPLENTWAPAPYSRSKLANKA VLAANGWNLKINGGALYTCALRPHY YIYGEGSRFLSVSINEALNNNGILSS VGKFSTVNPVYVGNVAWGHILAL RALQDPKKAPSIRGQFYYISDDTPH QSYDNLNYTLSKEFGPPPLDSRWS FPLSLMYWIGFLLGNR*GFLLRPHY TYRPFPRHISSHCSN*ALFHLLTRPHY TYRPFPRHISSHCSN*ALFHLTRPHY GREGLIGVLRPHY GREGSLVVYPWTQRPFDISFGNLSS ASAIMGKPPKSKAHGKVLTFLGT ACMWDPGGTFKLPGENVAGLTVFG QSHFRQKNFTPEGARFFLGRKMGD LELASALVPSRPLKPLGP QSHFRQKNFTPEGARFFLGRKMGD LELASALVPSRPLKPLGP QSHFRQKNFTPEGARFFLGRKMGD LELASALVPSRPLKPLGP GRAPPEDGGPLSLPNAMARGPKK HLRRVAAVPKHWMLDKLTGVFAPAR PSTGPHKLRECLPPIFFLRNRLKYA LTSDEVKKICMQRFIKIDGQVRTD LTSDEVKKICMQRFIKIDGQVRTD LTSDEVKKICMQRFIKIDGQVRTD LTSDEVKKICMQRFIKIDGQVRTD LTSDEVKKICMQRFIKIDGQVRTD LTSDEVKKICMQRFIKIDGQVRTD LTSDEVKICMQRFIKIDGQVRTD LTSDEVKICMQRFIKIDGQVRTD LTSDEVKICMQRFIKIDGQVRTD LTSDEVKICMQRFIKIDGQVRTD LTSDEVKICMQRFIKIDGQVRTD LTSDEVKICMQRFIKIDGQVRTD LTSDEVKICMQRFIKIDGQVRTD LTSDEVKICMQRFIKIDGQVRTD LTSDEVKICMQRFIKIDGQVRTD LTSDEVKICMAR TYTE					•	1		1
RSG\temples								
SCLTRDNGVAPAPAAMEAHPHFRE ERLEEINAGQQDCRRPGCSKQFPRIS RRKFPAAPPGSLRRQDKPRTTTKRE RRFPAAPPGSLRRQDKPRTTTKRE KTFLKVQGPSSGVCPQNKTQETPR 1748 7245 A 1878 87 260 1749 7246 A 1879 1 1254 1750 7247 A 1880 160 615 PSLNTYVTSPLSENFSARYRNHSND LTCVHTELQNKTKLTVLEGDILDEP FLKRACQ\DVSVNIHTACIIDVFGVI HRESIMNVNVKGRVAWGGDKARS GPEQKEGGEKRSLSIEHLLCSGP SDFADHYQLGELKAAIFSFIDEKTRT EQ SCYFGLATMGWSCLVTGAGGLLGC RIVKLLVEKELKEIRALDKAFREPL REEFISKLQ\NIKTKLTVLEGDILDEP LKESLARDRLRSIIHTACFHLMSFGV VITHEFFNMVQC*KVPSSC*EACVQ ASVPVFIYTSSIEVAGPNSYKEIIQNC ASVPVFIYTSSIEVAGPNSYKEIIQNC HEEPPLENTWPAPYPRSKKLAKKA VLAANGWNLKNIGGALYTCALRPW YTYGEGSRFLSVSINEALNNGILSS VGKFSTIVNPVYVGNVAWGHILAL RALQDPKKAFSIRGQFYYISDDTPH QSYDNLNYTLISKEGGPPPLDSRWS FPLSLMYW\GFLLGNR*GFLL\RPIY TYRPFPNRHISSHCSN*ALFHLLFIEGGFSEILGVLRRLITAGGGKAKAGKR VGSWVWVPFVDPAQGRNLEVPRIQ GFSEILGVLRRLITAGGGKAKAGKR VGSWVWVPFVDPAQGRNLEVPRIQ TYRPFFNRHISSHCSN*ALFHLLFIEGGFSEILGVLRRLITAGGGKAKAGKR VGSWVWVPFVDPAQGRNLEVPRIQ GKGSLVVYP\WTQRFVFDISFGNLSS ASAIMGKPPKSKAHGIKKVLTTC*Q ACMWDPGGTFKLPGENVAGLTVFG QSHFRQKNITPEGARFFLGRKMGD LELASALVPSRLPLKPTLGTPFLNRKKYC LTSUDEVKKICMQRFIKIDGQVRITD				1	ł	1		
				ł				
RPKFPAAPPGSLRRQDKPRNTITKRR KTFLKVQGPSSGVCPQNKTQETPR				1				
1747 7244 A 1877 1 1059				1		ł		ERLEEI\AGQQDCRRPGCSKQFPRIS
1747 7244 A 1877 1 1059					ļ	j		RFKFPAAPPGSLRRQDKPRF\TTKRP
1747	į			1	1	j .		
1748 7245 A 1878 87 260		1747	7244	+	1877	1	1050	1111 211 Q 01 00 0 1 Q 1 M 1 Q 2 1 M
1749 7246 A 1879 1 1254						·		<u> </u>
1750 7247 A 1880 160 615 PSLNTYVTSPLSENFSARYRNHSND LTCVHTELQNKTKLTVLEGDILDEP FLKRACQ\DVSV\\INTACIDVFGVI\\ HRESIMN\VN\KGRVAWGGDKARW GNEDQKEGQEGKRSLSIEHLLCSGP SDFADHYQLGELKAAIFSFIDEKTR\\ EQ CPLQGHPRVTLESDLLPSIFCFLVSD SCYFGLATMGWSCLVTGAGGLLGC RIVRLLVEEKELKEIRALDKAFRPEL REFSKLQ\\NK\\TKLTVLEGDILDEP LKESLARDRLRSIIHTACFHLMSFGV\\THREFF\MN\QC*KVPSSC*EACVQ AS\VPVFIYTSSIEVAGPNSYKEIIQNC HEEPLEN\TWPAPYPRSKKLA\KKAVLAANGWNLK\\NGGAL\TCALRPM YIYGEGSRFLS\SINEAL\NN\GILSS VGKFST\\NPYY\GNVAWGHLLAL RALQDPKKAPSIRG\GYY\ISDDTPH QS\TMYTL\SKE\FGPPPLDSR\\MSFFLS\MY\WIGFLLGNR*GFLL\\MPYY TYRPFF\NR\HISSHCS\N**AL\FLL\F\K\KY\MY\GFLL\\GRA\GK\KY\KY\FS\TMY\\GFL\\GRA\GK\KY\TFL\GF\K\GK\GK\KY\TFL\GT\\GT\K\GK\GK\KY\TFL\GT\\GT\K\GK\GK\KY\TFL\GT\\GK\GK\GK\KY\TFL\GT\\GK\GK\GK\KY\TF\\GK\GK\GK\KY\TFL\GT\\GK\GK\GK\KY\TFL\GT\\GK\GK\GK\KY\TFL\GT\\GK\GK\GK\KY\TFL\GT\\GK\GK\GK\FT\\GT\GK\GK\GK\KY\TFL\GT\\GK\GK\GT\K\TP\\GK\GT\K\GT\\GT\\GK\GK\GK\FT\\GT\GT\GT\\GK\GK\GK\FT\\GT\GT\GT\GT\\GT\\GT\\GT\\GT\\GT\\GT								
LTCVHTELQNKTKLTVLEGDILDEP FILKRACQIDVSVINIHTACIIDVFGVIT HRESIMNVNVKGRVAWGGDKARW GNEDQKEGQEGKRSLSIEHLLCSGP SDFADHYQLGELKAAIFSFIDEKTRT EQ 1751 7248 A 1881 53 1338 CPLQGHPRVTLESDLLPSIFCFLVSD SCYFGLATMGWSCLVTGAGGLLGC RIVRLLVEEKELKEIRALDKAFRPEL REEFSKLQNIKUTKLTVLEGDILDEP LKSELARDRLRSIHTACFHLMSFGV VTHREFF\MNVQC*KVPSSC*EACVQ ASVPVFIYTSSIEVAGPNSYKEIIQNC HEEEPLENTWPAPYPRSKLANKAG VLAANGWNLKINGGALYTCALRPM YIYGEGSRFLSVSINEALNNNGILSS VGKFSTLVNPYYVGNVAWGHILAL RALQDPKKAPSIRGGFYYISDDTPH QSYDNLNYTLSKEFGPPPLDSRWS FPLSLMYWVIGFLLGNR*GFLLVRPIY TYXRPFTNRHISSHCSN*ALFHLLFIKE GFSEILGVLRPLLTAGALFHLLFIKE GFSEILGVLRPLLTAGAGKAKAGKR VGSWVWVPFVDPAQGRNLEVPRIQ DIKATITISLWGKVVNVEDAGGENTP GKGSLVYPIWTQRFFDSFGNLSS ASAIMGKPPKSKAHGKKVLTFLGT MPTKHLE*FSRGTFCPSLK*TCTC*Q ACMWDPGGTFKLPGENVAGLTVFG QSHFRQKNFTPEGARFFLGRKMGD LELASALVPSRLPKPLGP 1753 7250 A 1883 1 960 GRPAPEDGGPLS.PNAAMARGPKK HLKRVAAIPKHWMLDLTGVFAR PSTGPHKLIRECLPFIIFLRNRLKYA LTS\DEVKKICMQRFIKI\DGQVR\TD						, -		
FLKRACQ\DVSV\IHTACIIDVFGVI\ HRESIMNVNVKGRVAWGGDKARW GNEDQKEGQEKKRSLSIEHLLCSGP SDFADHYQLGELKAAIFSFIDEKTRI EQ 1751 7248 A 1881 53 1338 CPLQGHPRVTLESDLLFSIFCFLVSD SCYFGLATMGWSCLVTGAGGLLGG RIVRLLVEEKELKEIRALDKAFRPEL REEFSKLQ\NK\TKLTVLEGDILDEPF LKESLARDRLRSIIHTACFHLMSFGV \(\text{VTHREFP\MNVQC**}\XYPSSC**PSC**ACVQ \(\text{ASVPVFIYTSSIEVAGPNSYKEIIQNG}\) HEEEPLENTWPAPYPRSKKLA\KKA \(\text{VLAANGWNLK\NGGALYTCALRPW}\) YIYGEGSRFLSVSINEALNN\GILSS \(\text{VGKFST\V\NP\YY\GNAWGHILAL}\) RALQDPKKAPSIRGQFYYISDDTPH \(\Q\text{QSYDNL\NYTL\SKE\FGP\PPLDSR\WS}\) FPLSLMYWIGFLLGNR*\GFLL\SIN\GF\SEL\ST\X\T\T\Y\T\Y\PPF\NR\HISSHCS\N**AL\FHLL\T\K\GF\SEL\SI\X\T\Y\T\Y\T\Y\T\Y\T\Y\T\Y\T\Y\T\Y\T\Y\T	1	1750	7247	A	1880	160	615	PSLNTYVTSPLSENFSARYRNHSND
HRESIMNVNVKGRVAWGGDKARW GNEDQKEGQEGKRSLSIEHILLCSGP SDFADHYQLGELKAAIFSFIDEKTRT EQ 1751 7248 A 1881 53 1338 CPLQGHPRVTLESDLLPSIFCFLVSD SCYFGLATMGWSCLVTGAGGLLGG RIVRLLVEEKELKEIRALDKAFRPEL REEFSKLQNK\TKLTVLEGDILDEP LKESLARDRLRSIIHTACFHLMSFGV \threfp\mnvQc*kvpssc*eacvQ ASVPVFIYTSSIEVAGPNSYKEIIQNG HEEPLENTWPAPYPSKKLAKKA VLAANGWNLKVNGGALYTCALRPM YIYGEGSRFLSVSINEALNNNGILSS VGKFST\VNPVYVGNVAWGHILAL RALQDPKKAPSIRGQFYYISDDTPH QSYDNLNYTLISKE\FGPPPLDSRWS FPLSLMYWIGFLLGNR*GFL\RPPY TYRPPFNRHISSHCSN*ALFHILFIKE GFSEILGVLRPLLTAGGGKAKAGKR VGSWVWPFVDPAQGRNLEVPRIQ 1752 7249 A 1882 3 575 HSLFGTSEVINKLLVPDAWGHFTEE DKATI\TSLWGK\VNVE\DAGGE\TP GKGSLVVYP\WTQRF\FD\SFGNLSS ASANMGRPKSKAHGIKKVLTFLGT MPTKHLE*FSRGTFCPSLK*TCTC*Q ACMWDPGGTTFKLPGENVAGLTVFG QSHFRQKNFTPEGARFFLGRKMGD LELASALVPSRLPLKPLGP 1753 7250 A 1883 1 960 GRPAPEDGGPLSLPNAAMARGPKK HLKRVAAIPKHWMLDKLTGVFAPR PSTGPHKL\RECLPFIIPLRNRLKYA LTS\DEVKKICMQRFIKI\DGQVR\TD	i	*				i		LTCVHTELQNKTKLTVLEGDILDEP
HRESIMNVNVKGRVAWGGDKARW GNEDQKEGQEGKRSLSIEHILLCSGP SDFADHYQLGELKAAIFSFIDEKTRT EQ 1751 7248 A 1881 53 1338 CPLQGHPRVTLESDLLPSIFCFLVSD SCYFGLATMGWSCLVTGAGGLLGG RIVRLLVEEKELKEIRALDKAFRPEL REEFSKLQNK\TKLTVLEGDILDEP LKESLARDRLRSIIHTACFHLMSFGV \threfp\mnvQc*kvpssc*eacvQ ASVPVFIYTSSIEVAGPNSYKEIIQNG HEEPLENTWPAPYPSKKLAKKA VLAANGWNLKVNGGALYTCALRPM YIYGEGSRFLSVSINEALNNNGILSS VGKFST\VNPVYVGNVAWGHILAL RALQDPKKAPSIRGQFYYISDDTPH QSYDNLNYTLISKE\FGPPPLDSRWS FPLSLMYWIGFLLGNR*GFL\RPPY TYRPPFNRHISSHCSN*ALFHILFIKE GFSEILGVLRPLLTAGGGKAKAGKR VGSWVWPFVDPAQGRNLEVPRIQ 1752 7249 A 1882 3 575 HSLFGTSEVINKLLVPDAWGHFTEE DKATI\TSLWGK\VNVE\DAGGE\TP GKGSLVVYP\WTQRF\FD\SFGNLSS ASANMGRPKSKAHGIKKVLTFLGT MPTKHLE*FSRGTFCPSLK*TCTC*Q ACMWDPGGTTFKLPGENVAGLTVFG QSHFRQKNFTPEGARFFLGRKMGD LELASALVPSRLPLKPLGP 1753 7250 A 1883 1 960 GRPAPEDGGPLSLPNAAMARGPKK HLKRVAAIPKHWMLDKLTGVFAPR PSTGPHKL\RECLPFIIPLRNRLKYA LTS\DEVKKICMQRFIKI\DGQVR\TD				1				FLKRACO\DVSVI\IHTACIIDVFGV\T
GNEDQKEGQEGKRSLSIEHLLCSGP SDFADHYQLGELKAAIFSFIDEKTRI EQ CPLQGHPRVTLESDLLPSIFCFLVSD SCYFGLATMGWSCLVTGAGGLLGC RIVRLLVEEKELKEIRALDKAFRPEL REEFSKLQ\NK\ITKLIVLEGDILDEPI LKESLARDRLRSIIHTACFHLMSFGV \text{VITHREFFMNVQC*KVPSSC*EACVQ} ASVPVFIYTSSIEVAGPNSYKEIIQNC HEEEPLENITWPAPYPRSKKLAKKA VLAANGWNLK\NGGALYTCALRPM YIYGEGSRFLSVSINEALNNNGILSS VGKFST\VNPVYV\GNVAWGHILAL RALQDPKKAPSIRGQFYYISDDTPH QSYDNLNYTL\SKE\FGPPPLDSRWS FPLSLMYW\GFLLGNR*GFLL\RPIY TYRPPFNRHISSHCSN*ALFHLL\IKE GFSEILGVLRPLLTAGGGKAKAGKR VGSWVWVPFVDPA\GRHEPVRIQ 1752 7249 A 1882 3 575 HSLFGTSEVINKLLVPDA\MGHFTEE D\KATITSLWGK\VNVEDAGGE\TP GKGSLVVYP\WTQRF\FD\SFGNLSS ASA\MGKPPKSKAHG\KKVLTFLGT MPTKHLE*FSRGTFCPSLK*TCTC*Q ACMWDPGGTFKL\PGENVAGLT\FGG QSHFRQKNFTPEGARFFLGRKMGD LELASALVPSRLPLKPLGP 1753 7250 A 1883 1 960 GRPAPEDGGPLSLPNAAMARGPKK HLKRVAA\PKHWMLDKLTGVFAPR PSTGPHKLLRECLPFIIF\LRNRLKYA LTS\DEVKKICMQRFIK\DGQVR\TD						. ,		
SDFADHYQLGELKAAIFSFIDEKTRI EQ	1							
EQ I751 7248 A I881 53 I338 CPLQGHPRVTLESDLLPSIFCFLVSD SCYFGLATMGWSCLVTGAGGLLGC RIVRLLVEKELKEIRALDKAFRPEL REEFSKLQNKVTKLTVLEGDILDEPF LKESLARDRLRSIIHTACFHLMSFGV VITHREFFMNVQC*KVPSSC*EACVQ ASVPVFIYTSSIEVAGPNSYKEIIQNG HEEPLENTWPAPYPRSKKLAKKA VLAANGWNLK\NGGALYTCALRPM YIYGEGSRFLSVSINEALNNNGILSS VGKFSTVNPVYVVGINVAWGHILAL RALQDPKKAPSIRGQFFYISDDTPH QSYDNLNYTLSKEUFGPPPLDSRWS FPLSLMYW\IGFLLGNR*GFLLRPIY TYRPPFNRHISSHCSN*ALFHILLTIKE GFSEILGVLRPLLTTAGGKAKAGKR VGSWVWPFVDPAQGRNLEVPRIQ GFSEILGVLRPLLTTAGGKAKAGKR VGSWVWPFVDPAQGRNLEVPRIQ GKGSLVVYPWTQRFFDISFGNLSS ASAIMGKPPKSKAHGIKKVLTFLGT MPTKHLE*FSRGTFCPSLK*TCTC*Q ACMWDPGGTFKLPGENVAGLTVFG QSHFRQKNFTPEGARFFLGRKMGD LELASALVPSRLPLKPLGP GRPAPEDGGPLSIPNAGMARGPKK HLKRVAAIPKHWMLDKLTGVFAPR PSTGPHKLIRECLPFIIFILRNRLKYA LTS\DEVKKICMQRFIK\DGQVR\TD	ı			1	·			
1751 7248 A 1881 53 1338 CPLQGHPRVTLESDLLPSIFCFLVSD SCYFGLATMGWSCLVTGAGGLLGG RIVRLLVEKELKEIRALDKAFRFEIL REEFSKLQINK;TKLTVLEGDILDEPF LKESLARDRLRSIIHTACFHLMSFGV \ THREFF\MNVQC*KVPSSC*EACVQ ASVPVFIYTSSIEVAGPNSYKEIIQNG HEEEPLENTWPAPYPRSKKLAKKA VLAANGWNLKWGGALYTCALRPM YIYGEGSRFLSVSINEALNNNGILSS VGKFST\NPVYV\GNVAWGHILAL RALQDPKKAPSIRGQFYYISDDTPH QSYDNLNYTLSKEFGPPPLDSRWS FPLSLMYWIGFLLGNR*GFLLRPIY TYRPFFNRHISSHCSN*ALFHLLFIKE GFSEILGVLRPLLTAGGGKAKAGKR VGSWVWVPFVDPAQGRNLEVPRIQ HSLFGTSEVINKLLVPDA\MGHFTEE D\MATITISLWGK\VNVEDAGGETP GKGSLVYYP\WTQRFFD\SFGNLSS ASANMGKPKSKAHG\KKVLTFLGT MPTKHLE*FSRGTFCPSLK*TCTC*Q ACMWDPGGTFKLPGENVAGLTVFG QSHFRQKNFTPEGARFFLGRKMGD LELASALVPSRLPLKPLGP GRPAPEDGGPLSLPNAAMARGPKK HLKRVAA\PKHWMLDKLTGVFAPR PSTGPHKL\RCVELPFIIFLRNRKKYA LTS\DEVKKICMQRFIK\DGQVR\TD]]		j		,
SCYFGLATMGWSCLVTGAGGLLGG RIVRLLVEEKELKEIRALDKAFRPEL REEFSKLQINK!TKLTVLEGDILDEPF LKESLARDRLRSIIHTACFHLMSFGV \text{VHREFF}MNVQC*KVPSSC*EACVQ} ASVPVFIYTSSIEVAGPNSYKEIIQNG HEEEPLEMTWPAPYPRSKKLAKKA VLAANGWNLKINGGALYTCALRPM YIYGEGSRFLSVSINEALNINGILSS VGKFST\VNPVY\GNVAWGHILAL RALQDPKKAPSIRGQFYYISDDTPH QSYDNLNYTLISKEFGPPPLDSRWS FPLSLMYWIGFLLGNR*GFLLRPIY TYRPPFNRHISSHCSN*ALFHLLFIKE GFSEILGVLRPLLTAGGGKAKAGKR VGSWVWPFVDPAQGRNLEVPRIQ 1752 7249 A 1882 3 575 HSLFGTSEVINKLLVPDA\MGHFTEE D\KATITSLWGK\VNVE\DAGGE\TP GKGSLVYP\WYORFFD\SFGNLSS ASANMGKPPKSKAHG\KKVLTFLGT MPTKHLE*FSRGTFCPSLK*TCTC*Q ACMWDPGGTFKLPGENVAGLTVFG QSHFRQKNFTPEGARFFLGRKMGD LELASALVPSRLPLKPLGP QSHFRQKNFTPEGARFFLGRKMGD LELASALVPSRLPLKPLGP PSTGPHKL\RECLPFIIF\LRNRLKYA LTS\DEVKKICMQRFIKI\DGQVR\TD	٠			 				
RIVRLLVEEKELKEIRALDKAFRPEL REEFSKLQ\NK\TKLTVLEGDILDEPF LKESLARDRLRSIIHTACFHLMSFGV \THREFF\MNVQ*KVPSSC*EACVQ ASVPVFIYTSSIEVAGPNSYKEIIQNG HEEEPLEN\TWPAPYPRSKKLAKKA VLAANGWNLK\NGGALYTCALRPM YIYGEGSRFLSVSINEALNNNGILSS VGKFST\VNPVYV\GNVAWGHILAL RALQDPKKAPSIRGQFYYISDDTPH QSYDNLNYTL\SKE\FGPPPLDSRW\S FPLSLMYW\IGFLLGNR*GFLL\RPIY TYRPPFNRHISSHCSN*ALFHLL\TIKE GFSEILGVLRPLLTAGGGKAKAGKR VGSWVWYPFVDPAQGRNLEVPRIQ 1752 7249 A 1882 3 575 HSLFGTSEVINKLLVPDA\MGHTTEE D\KATI\TSL\WGK\NVVE\DAGGE\TIFT GKGSLVVYP\WTQRF\FD\SFGNLSS ASA\MGKPPKSKAHG\KKVLTFLGT MPTKHLE*FSRGTFCPSLK*TCTC*Q ACMWDPGGTFKLPGENVAGLTVFG QSHFRQKNFTPEGARFFLGRKMGD LELASALVPSRLPLKPLGP 1753 7250 A 1883 1 960 GRPAPEDGGPLSLPNAAMARGPKK HLKRVAA\PKH\WMLDKLTGVFAPR PSTGPHKL\RECLPFIIF\LRNRLKYA LTS\DEVKKICMQRFIK\NDGQVR\TD	1	1751	7248	A	1881	53	1338	
REEFSKLQ\NK\TKLTVLEGDILDEPF LKESLARDRLRSIIHTACFHLMSFGV \threff\mnvQc*kvpssc*eacvQ ASVPVFIYTSSIEVAGPNSYKEIIQNG HEEEPLEN\TWPAPYPRSKKLA\KKA \tilde{VLAANGWNLK\NGGALYTCALRPM} YIYGEGSRFLSVSINEALNNNGILSS VGKFST\VNPVYV\GNVAWGHILAL RALQDPKKAPSIRGQFYYISDDTPH QSYDNLNYTL\SKE\FGPPPLDSRW\S FPLSLMYWIGFLLGNR*GFLL\RPIY TYRPFNRHISSHCSN*ALFHLL\TKE GFSEILGVLRPLLTAGGGKAKAGKR VGSWVVPFVDPAQGRNLEVPRIQ 1752 7249 A 1882 3 575 HSLFGTSEVINKLLVPDA\MGHFTEE D\KATI\TSL\WGK\VNVE\DAGGE\TP GKGSLVVYP\WTQRF\FD\SFGNLSS ASA\MGKPPKSKAHG\KKVLTFLGT MPTKHL\PFSRGTFCPSLK*TCTC*Q ACMWDPGGTFKLPGENVAGLTVFG QSHFRQKNFTPEGARFFLGRKMGD LELASALVPSRLPLKPLGP 1753 7250 A 1883 1 960 GRPAPEDGGPLSLPNAAMARGPKK HLKRVAA\PKH\WMLDKLTGVFAPR PSTGPHKL\RECLPFIIF\LRNRLKYA LTS\DEVKKICMQRFIK\NDGQVR\TD	-			1				SCYFGLATMGWSCLVTGAGGLLGQ
LKESLARDRLRSIIHTACFHLMSFGV \(\text{THREFF\MNVQC*KVPSSC*EACVQ}\) ASVPVFIYTSSIEVAGPNSYKEIIQNG HEEEPLENITWPAPYPRSKKLA\KKA \(\text{VLAANGWNLK\NGGALYTCALRPM}\) YIYGGGSRFLSVSINEALNNNGILSS \(\text{VGKFST\VNPYY\GNVAWGHILAL}\) RALQDPKKAPSIRGQFYYISDDTPH \(\QSYDNLNYTL\SKE\FGPPPLDSR\W\S\) \(\text{FPLSLMYWIGFLLGNR*GFLL\RPIY}\) \(\text{TYRPFFNRHISSHCSN*ALFHLLTIKE}\) \(\text{GFSEILGVLRPLLTAGGGKAKAGKR}\) \(\text{VGSWVWPFVDPAQGRNLEVPRIQ}\) \(\text{VGSWVWPFVDPAQGRNLEVPRIQ}\) \(\text{VGSWVWPFVDPAQGRNLEVPRIQ}\) \(\text{DIKATITSLWGK\VNVE\DAGGEVIP}\) \(\text{GKGSLVVYP\WTQRF\FD\SFGNLSS}\) \(\text{ASAI\MGKPPKSKAHG\KK\VLTFLGT}\) \(\text{MPTKHLE*FSRGTFCPSLK*TCTC*Q}\) \(\text{ACMWDPGGTFKLPGENVAGLTVFG}\) \(\text{QSHFRQKN\FTDEGAR\FFLGRK\MGD}\) \(\text{LELASALVPSRLPLK\PLGP}\) \(\text{LELASALVPSRLPLK\PLGP}\) \(\text{LELASALVPSRLPLK\PLGP}\) \(\text{PSTGPHKL\RECL\PFII\F\LR\RRLK\YA}\) \(\text{LTS\DEVKKICMQ\RF\IK\\DG\QVR\TD}\)								RIVRLLVEEKELKEIRALDKAFRPEL
LKESLARDRLRSIIHTACFHLMSFGV \(\text{THREFF\MNVQC*KVPSSC*EACVQ}\) ASVPVFIYTSSIEVAGPNSYKEIIQNG HEEEPLENITWPAPYPRSKKLA\KKA \(\text{VLAANGWNLK\NGGALYTCALRPM}\) YIYGGGSRFLSVSINEALNNNGILSS \(\text{VGKFST\VNPYY\GNVAWGHILAL}\) RALQDPKKAPSIRGQFYYISDDTPH \(\QSYDNLNYTL\SKE\FGPPPLDSR\W\S\) \(\text{FPLSLMYWIGFLLGNR*GFLL\RPIY}\) \(\text{TYRPFFNRHISSHCSN*ALFHLLTIKE}\) \(\text{GFSEILGVLRPLLTAGGGKAKAGKR}\) \(\text{VGSWVWPFVDPAQGRNLEVPRIQ}\) \(\text{VGSWVWPFVDPAQGRNLEVPRIQ}\) \(\text{VGSWVWPFVDPAQGRNLEVPRIQ}\) \(\text{DIKATITSLWGK\VNVE\DAGGEVIP}\) \(\text{GKGSLVVYP\WTQRF\FD\SFGNLSS}\) \(\text{ASAI\MGKPPKSKAHG\KK\VLTFLGT}\) \(\text{MPTKHLE*FSRGTFCPSLK*TCTC*Q}\) \(\text{ACMWDPGGTFKLPGENVAGLTVFG}\) \(\text{QSHFRQKN\FTDEGAR\FFLGRK\MGD}\) \(\text{LELASALVPSRLPLK\PLGP}\) \(\text{LELASALVPSRLPLK\PLGP}\) \(\text{LELASALVPSRLPLK\PLGP}\) \(\text{PSTGPHKL\RECL\PFII\F\LR\RRLK\YA}\) \(\text{LTS\DEVKKICMQ\RF\IK\\DG\QVR\TD}\)	١			1 1				REEFSKLO\NK\TKLTVLEGDILDEPF
Cheffymnvqc*kvpssc*eacvq asvpvfiytssievagpnsykeiiqno heeeplen/twpapyprskklakka vlaangwnlkinggalytcalrpm yiygegsrflsvsinealmnngilss vgkfstivnpvyv/gnvawghilal ralqdpkkapsirgofyyisddtph qsydnlnytliskeigpppldsrws fplslmywigfllgni*gfllirpikeigfseilgvlrplltagggkakagkr vgswvwpfvdpaqgrnlevpriq dfseilgvlrplltagggkakagkr vgswvwvpfvdpaqgrnlevpriq dkativtslwgk/vnvedaggeitp dkativtslwgk/vnvedaggeitp gkgslvvypiwtqrfpisfgnlss asaingkppkskahgikkvltflgt mptkhle*fsrgtfcpslk*ttcteq acmwdpggtfklpgenvagltvfg qshfrqknftpegarfflgrkmgd lelasalvpsrlplkplgp 1753 7250 A 1883 1 960 Grpapedggplslpnaamargpkk hlkrvaaipkhwmldkltgvfapr pstgphkl/reclpfiifilrnrlkya lts/devkkicmqrfiki/dgqvr/td	-			1 1				
ASVPVFIYTSSIEVAGPNSYKEIIQNO HEEEPLEN\TWPAPYPRSKKLA\KKA VLAANGWNLK\NGGALYTCALRPM YIYGEGSRFLSVSINEALNNNGILSS VGKFST\VNPVY\VGNVAWGHILAL RALQDPKKAPSIRGOFYYISDDTPH QSYDNLNYTL\SKE\FGPPPLDSRW\S FPLSLMYW\IGFLLGNR*GFLL\RPIY TYRPFNRHISSHCSN*ALFHLL\PIKE GFSEILGVLRPLLTAGGGKAKAGKR VGSWVWVPFVDPAQGRNLEVPRIQ 1752 7249 A 1882 3 575 HSLFGTSEVINKLLVPDA\MGHFTEE D\KATI\TSLWGK\VNVE\D\AGGE\TP GKGSLVYP\WTQRF\FD\SFGNLSS ASA\MGKPPKSKAHG\KKVLTFLGT MPTKHLE*FSRGTFCPSLK*TCTC*Q ACMWDPGGTFKLPGENVAGLTVFG QSHFRQKNFTPEGARFFLGRKMGD LELASALVPSRLPLKPLGP 1753 7250 A 1883 I 960 GRPAPEDGGPLSLPNAAMARGPKK HLKRVAA\PKHWMLDKLTGVFAPR PSTGPHKL\RECLPFIIF\LRNRLKYA LTS\DEVKKICMQRFIKI\DGQVR\TD		i		1 1	·			
HEEEPLEN\TWPAPYPRSKKLA\KKA VLAANGWNLK\NGGALYTCALRPM YIYGEGSRFLSVSINEALNNNGILSS VGKFST\VNPVYV\GNVAWGHILAL RALQDPKKAPSIRG\(\text{GFYYISDDTPH}\) QSYDNLNYTL\SKE\(\text{FGPPPLDSRWS}\) FPLSLMYW\(\text{IGFNR}\) TYRPPFNRHISSHCSN*ALFHLL\(\text{FIKE}\) GFSEILGVLRPILTAGGGKAKAGKE VGSWVWVPFVDPAQGRNLEVPRIQ PSSWVWVPFVDPAQGRNLEVPRIQ UKATITISLW\(\text{GK}\)\VNVE\D\(\text{AGGENTP}\) GKGSLVVYP\WTQR\(\text{FD}\)\SGGE\(\text{TP}\) GKGSLVVYP\WTQR\(\text{FD}\)\SGGE\(\text{TP}\) MPTKHLE*FSRGTFCPSLK*TCTC*Q ACMWDPGGTFKLPGENVAGLTVFG QSHFRQKNFTPEGARFFLGRKMGD LELASALVPSRLPLKPLGP 1753 7250 A 1883 1 960 GRPAPEDGGPLSLPNAAMARGPKK HLKRVAA\PKHWMLDKLTGVFAPR PSTGPHKL\(\text{RECLPFIIF\}\)LRNRLKYA LTS\(\text{DEVKKICMQRFIKI\DGQVR\(\text{TD}\)	J							
VLAANGWNLK\NGGALYTCALRPM YIYGEGSRFLSVSINEALNNNGILSS VGKFST\VNPVYV\GNVAWGHILAL RALQDPKKAPSIRGQFYYISDDTPH QSYDNLNYTL\SKE\FGPPPLDSRW\S FPLSLMYW\IGFLLGNR*GFLL\RPIY TYRPPFNRHISSHCSN*ALFHLLFIKE GFSEILGVLRPLLTAGGGKAKAGKR VGSWVWVPFVDPAQGRNLEVPRIQ 1752 7249 A 1882 3 575 HSLFGTSEVINKLLVPDA\MGHFTEE D\KATI\TSLWGK\VNVE\D\AGGE\TP GKGSLVVYP\WTQRF\FD\SFGNLSS ASAI\MGKPPKSKAHG\KKVLTFLGT MPTKHLE*FSRGTFCPSLK*TCTC*Q ACMWDPGGTFKLPGENVAGLTVFG QSHFRQKNFTPEGARFFLGRKMGD LELASALVPSRLPLKPLGP 1753 7250 A 1883 1 960 GRPAPEDGGPLSLPNAAMARGPKK HLKRVAA\PKHWMLDKLTGVFAPR PSTGPHKL\RECLPFIIF\LRNRLKYA LTS\DEVKKICMQRFIKI\DGQVR\TD	ı	ł				i		
YIYGEGSRFLSVSINEALNINGILSS VGKFST\VNPVYV\GNVAWGHILAL RALQDPKKAPSIRGQFYYISDDTPH QSYDNLNYTL\SKE\FGPPPLDSRWS FPLSLMYW\IGFLLGNR*GFLL\RPIY TYRPPFNRHISSHCSN*ALFHLL\FIKE GFSEILGVLRPLLTAGGGKAKAGKR VGSWVWVPFVDPAQGRNLEVPRIQ 1752 7249 A 1882 3 575 HSLFGTSEVINKLLVPDA\MGHFTEE D\KATI\TSLWGK\VNVE\DAGGE\TP GKGSLVVYP\WTQRF\FD\SFGNLSS ASA\MGKPPKSKAHG\KKVLTFLGT MPTKHLE*FSRGTFCPSLK*TCTC*Q ACMWDPGGTFKLPGENVAGLTVFG QSHFRQKNFTPEGARFFLGRKMGD LELASALVPSRLPLKPLGP 1753 7250 A 1883 1 960 GRPAPEDGGPLSLPNAAMARGPKK HLKRVAA\PKHWMLDKLTGVFAPR PSTGPHKL\RECLPFIIF\LRNRLKYA LTS\DEVKKICMQRFIKI\DGQVR\TD	i						·	
VGKFST\VNPVYV\GNVAWGHILAL RALQDPKKAPSIRGQFYYISDDTPH QSYDNLNYTL\SKE\FGPPPLDSRWS FPLSLMYW\IGFLLGNR*GFLL\RPIY TYRPPFNRHISSHCSN*ALFHLL\FIKE GFSEILGVLRPLLTAGGGKAKAGKR VGSWVWVPFVDPAQGRNLEVPRIQ 1752 7249 A 1882 3 575 HSLFGTSEVINKLLVPDA\MGHFTEE D\KATI\TSLWGK\VNVE\DAGGE\TP GKGSLVVYP\WTQRF\FD\SFGNLSS ASA\MGKPPKSKAHG\KKVLTFLGT MPTKHLE*FSRGTFCPSLK*TCTC*Q ACMWDPGGTFKLPGENVAGLTVFG QSHFRQKNFTPEGARFFLGRKMGD LELASALVPSRLPLKPLGP 1753 7250 A 1883 1 960 GRPAPEDGGPLSLPNAAMARGPKK HLKRVAA\PKHWMLDKLTGVFAPR PSTGPHKL\RECLPFIIF\LRNRLKYA LTS\DEVKKICMQRFIKI\DGQVR\TD	1	1		1 1	,	1	l	
RALQDPKKAPSIRGQFYYISDDTPH QSYDNLNYTL\SKE\FGPPPLDSRWS FPLSLMYW\IGFLLGNR*GFLL\RPIY TYRPPFNRHISSHCSN*ALFHLL\FIKE GFSEILGVLRPLLTAGGGKAKAGKR VGSWVWVPFVDPAQGRNLEVPRIQ 1752 7249 A 1882 3 575 HSLFGTSEVINKLLVPDA\MGHFTEE D\KATI\TSLWGK\VNVE\DAGGE\TP GKGSLVVYP\WTQRF\FD\SFGNLSS ASA\MGKPPKSKAHG\KKVLTFLGT MPTKHLE*FSRGTFCPSLK*TCTC*Q ACMWDPGGTFKLPGENVAGLTVFG QSHFRQKNFTPEGARFFLGRKMGD LELASALVPSRLPLKPLGP 1753 7250 A 1883 1 960 GRPAPEDGGPLSLPNAAMARGPKK HLKRVAA\PKHWMLDKLTGVFAPR PSTGPHKL\RECLPFIIF\LRNRLKYA LTS\DEVKKICMQRFIKI\DGQVR\TD					,]	ļ	
RALQDPKKAPSIRGQFYYISDDTPH QSYDNLNYTL\SKE\FGPPPLDSRWS FPLSLMYW\IGFLLGNR*GFLL\RPIY TYRPPFNRHISSHCSN*ALFHLL\FIKE GFSEILGVLRPLLTAGGGKAKAGKR VGSWVWVPFVDPAQGRNLEVPRIQ 1752 7249 A 1882 3 575 HSLFGTSEVINKLLVPDA\MGHFTEE D\KATI\TSLWGK\VNVE\DAGGE\TP GKGSLVVYP\WTQRF\FD\SFGNLSS ASA\MGKPPKSKAHG\KKVLTFLGT MPTKHLE*FSRGTFCPSLK*TCTC*Q ACMWDPGGTFKLPGENVAGLTVFG QSHFRQKNFTPEGARFFLGRKMGD LELASALVPSRLPLKPLGP 1753 7250 A 1883 1 960 GRPAPEDGGPLSLPNAAMARGPKK HLKRVAA\PKHWMLDKLTGVFAPR PSTGPHKL\RECLPFIIF\LRNRLKYA LTS\DEVKKICMQRFIKI\DGQVR\TD							I	VGKFST\VNPVYV\GNVAWGHILAL
QSYDNLNYTL\SKEFGPPPLDSRWS FPLSLMYW\IGFLLGNR*GFLL\RPIY TYRPPFNRHISSHCSN*ALFHLLFIKE GFSEILGVLRPLLTAGGGKAKAGKR VGSWVWVPFVDPAQGRNLEVPRIQ 1752 7249 A 1882 3 575 HSLFGTSEVINKLLVPDA\MGHFTEE D\KATI\TSLWGK\VNVE\DAGGE\TP GKGSLVVYP\WTQRF\FD\SFGNLSS ASA\MGKPPKSKAHG\KKVLTFLGT MPTKHLE*FSRGTFCPSLK*TCTC*Q ACMWDPGGTFKLPGENVAGLTVFG QSHFRQKNFTPEGARFFLGRKMGD LELASALVPSRLPLKPLGP 1753 7250 A 1883 1 960 GRPAPEDGGPLSLPNAAMARGPKK HLKRVAA\PKHWMLDKLTGVFAPR PSTGPHKL\RECLPFIIF\LRNRLKYA LTS\DEVKKICMQRFIKI\DGQVR\TD	1				. [. [
FPLSLMYW\IGFLLGNR*GFLL\RPIY TYRPPFNRHISSHCSN*ALFHLLFIKE GFSEILGVLRPLLTAGGGKAKAGKR VGSWVWVPFVDPAQGRNLEVPRIQ 1752 7249 A 1882 3 575 HSLFGTSEVINKLLVPDA\MGHFTEE D\KATI\TSLWGK\VNVE\DAGGE\TP GKGSLVVYP\WTQRF\FD\SFGNLSS ASA\MGKPPKSKAHG\KKVLTFLGT MPTKHLE*FSRGTFCPSLK*TCTC*Q ACMWDPGGTFKLPGENVAGLTVFG QSHFRQKNFTPEGARFFLGRKMGD LELASALVPSRLPLKPLGP 1753 7250 A 1883 1 960 GRPAPEDGGPLSLPNAAMARGPKK HLKRVAA\PKHWMLDKLTGVFAPR PSTGPHKL\RECLPFIIF\LRNRLKYA LTS\DEVKKICMQRFIKI\DGQVR\TD	ł	. 1				· }	ł	
TYRPPFNRHISSHCSN*ALFHLLFIKE GFSEILGVLRPLLTAGGGKAKAGKR VGSWVWVPFVDPAQGRNLEVPRIQ 1752 7249 A 1882 3 575 HSLFGTSEVINKLLVPDA\MGHFTEE D\KATI\TSLWGK\VNVE\DAGGE\TP GKGSLVVYP\WTQRF\FD\SFGNLSS ASA\MGKPPKSKAHG\KKVLTFLGT MPTKHLE*FSRGTFCPSLK*TCTC*Q ACMWDPGGTFKLPGENVAGLTVFG QSHFRQKNFTPEGARFFLGRKMGD LELASALVPSRLPLKPLGP 1753 7250 A 1883 1 960 GRPAPEDGGPLSLPNAAMARGPKK HLKRVAA\PKHWMLDKLTGVFAPR PSTGPHKL\RECLPFIIF\LRNRLKYA LTS\DEVKKICMQRFIKI\DGQVR\TD	I	· 1					·	
GFSEILGVLRPLLTAGGGKAKAGKR VGSWVWVPFVDPAQGRNLEVPRIQ 1752 7249 A 1882 3 575 HSLFGTSEVINKLLVPDA\MGHFTEE D\KATI\TSLWGK\VNVE\DAGGE\TP GKGSLVVYP\WTQRF\FD\SFGNLSS ASAI\MGKPPKSKAHG\KKVLTFLGT MPTKHLE*FSRGTFCPSLK*TCTC*Q ACMWDPGGTFKLPGENVAGLTVFG QSHFRQKNFTPEGARFFLGRKMGD LELASALVPSRLPLKPLGP 1753 7250 A 1883 1 960 GRPAPEDGGPLSLPNAAMARGPKK HLKRVAA\PKHWMLDKLTGVFAPR PSTGPHKL\RECLPFIIF\LRNRLKYA LTS\DEVKKICMQRFIKI\DGQVR\TD	١	ļ		1 1			ļ	
1752 7249 A 1882 3 575 HSLFGTSEVINKLLVPDA\MGHFTEE D\KATI\TSLWGK\VNVE\DAGGE\TP GKGSLVVYP\WTQRF\FD\SFGNLSS ASA\MGKPPKSKAHG\KKVLTFLGT MPTKHLE*FSRGTFCPSLK*TCTC*Q ACMWDPGGTFKLPGENVAGLTVFG QSHFRQKNFTPEGARFFLGRKMGD LELASALVPSRLPLKPLGP 1753 7250 A 1883 1 960 GRPAPEDGGPLSLPNAAMARGPKK HLKRVAA\PKHWMLDKLTGVFAPR PSTGPHKL\RECLPFIIF\LRNRLKYA LTS\DEVKKICMQRFIKI\DGQVR\TD	ĺ	l			Ì	- 1		
1752 7249 A 1882 3 575 HSLFGTSEVINKLLVPDA\MGHFTEE D\KATI\TSLWGK\VNVE\DAGGE\TP GKGSLVVYP\WTQRF\FD\\SFGNLSS ASAI\MGKPPKSKAHG\KKVLTFLGT MPTKHLE*FSRGTFCPSLK*TCTC*Q ACMWDPGGTFKLPGENVAGLTVFG QSHFRQKNFTPEGARFFLGRKMGD LELASALVPSRLPLKPLGP 1753 7250 A 1883 1 960 GRPAPEDGGPLSLPNAAMARGPKK HLKRVAA\PKHWMLDKLTGVFAPR PSTGPHKL\RECLPFIIF\LRNRLKYA LTS\DEVKKICMQRFIKI\DGQVR\TD	I	1			ļ	. }	ł	
D\KATI\TSLWGK\VNVE\DAGGE\TP GKGSLVVYP\WTQRF\FD\SFGNLSS ASAI\MGKPPKSKAHG\KKVLTFLGT MPTKHLE*FSRGTFCPSLK*TCTC*Q ACMWDPGGTFKLPGENVAGLTVFG QSHFRQKNFTPEGARFFLGRKMGD LELASALVPSRLPLKPLGP 1753 7250 A 1883 I 960 GRPAPEDGGPLSLPNAAMARGPKK HLKRVAA\PKHWMLDKLTGVFAPR PSTGPHKL\RECLPFIIF\LRNRLKYA LTS\DEVKKICMQRFIKI\DGQVR\TD	ļ			Ш				
GKGSLVVYP\WTQRF\FD\SFGNLSS ASAI\MGKPPKSKAHG\KKVLTFLGT MPTKHLE*FSRGTFCPSLK*TCTC*Q ACMWDPGGTFKLPGENVAGLTVFG QSHFRQKNFTPEGARFFLGRKMGD LELASALVPSRLPLKPLGP 1753 7250 A 1883 1 960 GRPAPEDGGPLSLPNAAMARGPKK HLKRVAA\PKHWMLDKLTGVFAPR PSTGPHKL\RECLPFIIF\LRNRLKYA LTS\DEVKKICMQRFIKI\DGQVR\TD	1	1752	7249	A	1882	3	575	HSLFGTSEVINKLLVPDA\MGHFTEE
GKGSLVVYP\WTQRF\FD\SFGNLSS ASAI\MGKPPKSKAHG\KKVLTFLGT MPTKHLE*FSRGTFCPSLK*TCTC*Q ACMWDPGGTFKLPGENVAGLTVFG QSHFRQKNFTPEGARFFLGRKMGD LELASALVPSRLPLKPLGP 1753 7250 A 1883 1 960 GRPAPEDGGPLSLPNAAMARGPKK HLKRVAA\PKHWMLDKLTGVFAPR PSTGPHKL\RECLPFIIF\LRNRLKYA LTS\DEVKKICMQRFIKI\DGQVR\TD	۱	[1	1	. 1	1	D\KATI\TSLWGK\VNVE\DAGGE\TP
ASAI\MGKPPKSKAHG\KKVLTFLGT MPTKHLE*FSRGTFCPSLK*TCTC*Q ACMWDPGGTFKLPGENVAGLTVFG QSHFRQKNFTPEGARFFLGRKMGD LELASALVPSRLPLKPLGP 1753 7250 A 1883 1 960 GRPAPEDGGPLSLPNAAMARGPKK HLKRVAA\PKHWMLDKLTGVFAPR PSTGPHKL\RECLPFIIF\LRNRLKYA LTS\DEVKKICMQRFIKI\DGQVR\TD	1	· I					1	
MPTKHLE*FSRGTFCPSLK*TCTC*Q ACMWDPGGTFKLPGENVAGLTVFG QSHFRQKNFTPEGARFFLGRKMGD LELASALVPSRLPLKPLGP 1753 7250 A 1883 1 960 GRPAPEDGGPLSLPNAAMARGPKK HLKRVAA\PKHWMLDKLTGVFAPR PSTGPHKL\RECLPFIIF\LRNRLKYA LTS\DEVKKICMQRFIKI\DGQVR\TD	۱	i i			ļ	1	1	
ACMWDPGGTFKLPGENVAGLTVFG QSHFRQKNFTPEGARFFLGRKMGD LELASALVPSRLPLKPLGP 1753 7250 A 1883 1 960 GRPAPEDGGPLSLPNAAMARGPKK HLKRVAA\PKHWMLDKLTGVFAPR PSTGPHKL\RECLPFIIF\LRNRLKYA LTS\DEVKKICMQRFIKI\DGQVR\TD	ı	·	•		· 1	i	, [
QSHFRQKNFTPEGARFFLGRKMGD LELASALVPSRLPLKPLGP 1753 7250 A 1883 1 960 GRPAPEDGGPLSLPNAAMARGPKK HLKRVAA\PKHWMLDKLTGVFAPR PSTGPHKL\RECLPFIIF\LRNRLKYA LTS\DEVKKICMQRFIKI\DGQVR\TD	١	ł		11	1	1	ł	
LELASALVPSRLPLKPLGP 1753 7250 A 1883 1 960 GRPAPEDGGPLSLPNAAMARGPKK HLKRVAA\PKHWMLDKLTGVFAPR PSTGPHKL\RECLPFIIF\LRNRLKYA LTS\DEVKKICMQRFIKI\DGQVR\TD	ł	ł					1	
1753 7250 A 1883 I 960 GRPAPEDGGPLSLPNAAMARGPKK HLKRVAA\PKHWMLDKLTGVFAPR PSTGPHKL\RECLPFIIF\LRNRLKYA LTS\DEVKKICMQRFIKI\DGQVR\TD	J	.]		-		j	-}	
HLKRVAA\PKHWMLDKLTGVFAPR PSTGPHKL\RECLPFIIF\LRNRLKYA LTS\DEVKKICMQRFIKI\DGQVR\TD	l				[LELASALVPSRLPLKPLGP
HLKRVAA\PKHWMLDKLTGVFAPR PSTGPHKL\RECLPFIIF\LRNRLKYA LTS\DEVKKICMQRFIKI\DGQVR\TD	ſ	1753	7250	A	1883	1	960	GRPAPEDGGPLSLPNAAMARGPKK
PSTGPHKL\RECLPFIIF\LRNRLKYA LTS\DEVKKICMQRFIKI\DGQVR\TD	1			[1	,	3
LTS\DEVKKICMQRFIKI\DGQVR\TD	ļ	1			ļ	·	ŀ	1
	I	[1	·	f	
	İ				l	- 1	į	
ITYP\AGFMDVI\SIDKDGREFSVL/Y	١	. [1		
LIDTQGVRFCL*HRITP*GRAKVQSC	I	I	ĺ		ł	ļ		
I I I AKMRKILLWAPKGIPSSWVT\HDAR	L	I					[AKMRKILLWAPKGIPSSWVT\HDAR

SEQ ID	SEQ ID	Me	SEO ID NO	: Nucleotide	Nucleotide	Amino acid sequence (X=Unknown; *=Stop
NO: of	NO: of	tho	in USSN	location of	location of last	codon; /=possible nucleotide deletion; \=possible
nucleo-tide	peptide sequence	d.	09/770,160	first codon	codon for last	nucleotide insertion)
Judgacasee	Jedaciie.			sequence	peptide	
<u> </u>			<u> </u>		sequence	
		Ì	ł			NHPATPDPPSSKVN*YHFRLDLETG
		1	l	ł	1	KDYLISSKFDTW*PCVMVT\GGA\N
		1	İ			LGRNWVLITN\RERHPGIF*PLVHVK
1		1	ļ		1	\DANGNKLLATSDFSNIFWLLGKGN
		1			1	KPW\ISL\PRGKGIPPHHLLEERDKRL
1754	7251	A	1884	1	1218	AAKQSSWVKWGPWVTWSDLLVP
1134	1231	^	1004	1	1210	FFQNSARGAGAGWQLPWTRFVWT\
				f .		SGLLEINE\TLVIQQRGVRIYDGEEKI KFDAGTLLLSTHRLIWRDQKNHEC
		1				CMAILLSQI\VFIEEQA\AGIGKSAKI
l	1	-	1	1	1	VVHL\HPAPPNKEPGP\FQSSKNSYI
ł		1		1	ļ	KLSFKEHGQIEFYRRLSEEMTQRRW
1						ENMPVSQSLQTNRGPQPGRIRAVGI
		1				VGTERKLEEKRKETDKNISEAFEDL
	ŀ			}		SKLMIKAKEMVELSKSIANKIKDKO
	1		1.			GDITEDETIRFKSYL\LSMGIANPVT
i			1	1		RETYGSGTQYHM\QLAKQL\AWNIA
1			[RVPLEERGGIMSLTEVYCLVNRARG
ł	₹ `				ľ	MELLSPEDLVNACKMLEALKLPLR
	1	1	}		l	LRVFDSGVMVIELQSHKEEEMVAS
1 .]	ALETVSEMGSLTS*EFAKLVGMSVL
					İ	LAKERLLLAEKMGHLCRDDSVEGL
1755	7252	10	1885	179	361	RFYPNLFMTQS
1,755	1232		1005	173	301	MPKVCFVHNFLKTSSERDLFALMN TVGKKHSIMSEKGRSKKFLHLIDSK
	1			1		KNEDPHLDGTL*
1756	7253	A	1886	2	913	RRLLLFGWARSGAVSLGSAGVSSS
	1		,			GFLTAPHSRRLTAAAAAAGGAWRF
	ĺ	1 1		1		EAERHRGWGAEEEQQPEGGAVCPG
				ļ		TERPCAMAYAYLFKYIIIGDTGGGR\
	ł					SCLLLQFTDKRFQPSAMTLTNGVEF
		1 1				GARMITIDGKQIK\LQIW\DTAGQES\
				j		FRSITR\SYY\RGAAGALLVYDITR\R
	ļ					DTSTHLTTW/LEDA/RQHSHFQHGS
		1			•	LCLLGNKSDL\ESRKE/VSKKRKEGE
		1 1				SFLQPRNHGLHLPWKTSCKNCFPM*
	•	1 1		í í		KEAFINTSKRNFIEKIQ\EGVFDINNE A\NGIKIGP\QHAATNATHAG\NQGG
•		11		1		QQAGGGCC
1757	7254	A	1893	138	426	FIHSHCCIVFRLFIHFSLHPKVIHSPIN
	'	11		.		SLLRIFQF*AIMNSTV*NILIHVFW*V
	ł			1 1		YTFPF\GINPKKGIARL*GVYIFSFSIY
						CQTVFQSDCKKAPF
1758	7255	A	1894	45	1057	FLVFLVETGFHHVAQAVLELLASSD
1] [PPALAPPKCWDYRCELLRLAEFCFL
					ſ	RTEFWYLLFFFFWRRSLALSPRLEC
				1 1		SGANL\THCNLR/LPGFKQFSCLSLSS
		[1		SWDYRCMPPHLATFFVF/SVETGFH
						RVAQASLELLSSGSLPALA/FPKC\W
						DYRAKATV/WPSPGVSSFILGL*TS*
,				·		FHSLEPYLHAWKTTSHLPTKEALT
ļ			•		1	W/VSHTAKTKHLWILVSILMEF*VA LIS/SFFLGPGGK*T*VTAPQCPSLGQ
j			,) ·	,	DTLS*FLHAACTRSVPYPGLA/CGPS
					[LWLTRVLLLPTPP*QQHNP/DTLEKT
-				[·].		SFPGPHWIL*/TPQPSLSETPAPKVPP
					· .]	FPAFGSIPTHEEPGLP
1759	7256	A	1895	2	289	

SEQ ID NO: of nucleo-tide sequence	SEQ ID NO: of peptide sequence	tho	SEQ ID NO: in USSN 09/770,160	location of first codon for peptide sequence	peptide sequence	Amino acid sequence (X=Unknown; *=Stop codon; /=possible nucleotide deletion; \=possible nucleotide insertion)
1760	7257	A	1896	1	397	
1761	7258	A	1897	1	410	STMISPVLILFSSFLCHVAIAGRTCPK PDDLPFSTVVPLKTFYEPGEEITYSC KPGYVSRGGM/RKFICPLTGLWPIN TLKCTPRVCPFAGNLRKMGAVRLIT DFLNYSPTRFSFSLLTWGFILEWAL DS\AKCIEGG
1762	7259	A	1898	19	1215	CQCDSSTMIFSRCSSLFSSFLCHVAI AGRTCPKPDDLPFSTVVPLKTFYEP GUEITYSCKPGYVSRGGIEESLSCPL VTGTVGPFNTSGNVTPRVCPF\AGIFR KMGGRTLITTF*NYPNTDPVFSLLTL GF*FWNGALDFWPSCTGGKGKW\S P\ELPGLVAPII\CPP\PSIP\TGFATLH VLLRPFRLGNNSPPIGDTAVFECLA HNMAMFG\NDTIT\CTTHGKLDLNY PECRGSKMPPFPHQDPDNGIW*TYP CQNPNTLFTRVKAPHLGLPHDGIFS GMGPRKE\LEC*PQTWGKPGSWPLA PSW*KPSLVKGTPVKKRPTVV/YPQ GERVKDSREKFKEWECLHG**KFLS FCKNKEKKCSYTEDAQCIDGTIEVP
		٠,	1000	-	1446	KCFK\EHSSLAFWKT\DAS\DVKPC
1763	7260	$\frac{A}{A}$	1899	58	954	MGEVSGTSDCTDDQCRQVKKALEG
1764	7261	A				GKAARGHRSKIKIRFFRPGGLGPGP AITAVAGMPRVYIGRLSYQAREHA VERLLNGHAKILEVDLKNGYGFVE FDDLRDADDAVYELNGKDLCGERV IVEHARGPRRDGSYGSGRSGYGYR RSGRDKYGPPTRTEDRLIVEN\LTSR CSWQDLKDYMRQAGEVTYADAHK GRQKMKGVIEFVSYSDMKRALEKL DGTEVNGRKIRLVEDKPGSRRRRSY SRSR\SHSRSRSRSRSRSKSRSRSGSS KSSHSKSRSRSRSRSKSRSRSQSRSRSKSRSRSQ SRSRSKKEKSRSPSKDKS\RSRSHSA\ GKSRSKSKDQAE\EKFQNNDNV\GK PKSRSPSRHKSKSKSRSQERRVEE GRKRGSF*QG\/EAQEKSLRQSRSN\ SRSKAGSR*PVDRSRSKSKDKRKSR KRSREESRSRSRSRSKSERSRKRG\S KRDSKAS\SCKKKKKEDTDRSQSRS PSRSV\SKEREHA\RSLESSQREGRG ESENAGTNQEDPGPGPRSN\SKSKP NLPIRMHRSKIKSQASKTPISGPMSR SR\SASRSP\SRSRSKSRSRSQSRSRS KKEKSRSPSKDKSLQPQP
1765	7262	A	1901	3	180	
1766	7263	A		227	440	GMHNVCYVAVNE*FCGFIIR*SLAE RRQIS*EFQLFKFTLCLELILARRAC RESMA\$PVAGSWSHFPEREF
1767	7264	A	1903	2	438	HEELDTSERKIEFDSASGTYTLYLIN GDAHFEEPQSLWNVADLVHQSPPE EKAPLDLSCPQNLFTPK\QEIQWIRI GA\NVS\NFTFAP\STIIFH\LGHA\AM LGLMYVYWTQLNMF\QTLKYLAIL GSVTFLAGNRMLAQQAVKRTAH



SEQ ID NO: of nucleo-tide sequence	SEQ ID NO: of peptide sequence	tho	SEQ ID NO: in USSN · 09/770,160	Nucleotide location of first codon for peptide sequence	Nucleotide location of last codon for last amino acid of peptide sequence	Amino acid sequence (X=Unknown; *=Stop codon; /=possible nucleotide deletion; \=possible nucleotide insertion)
1760	7257	Α	1896	1	397.	
1761	7258	A	1897	1	410	STMISPVLILFSSFLCHVAIAGRTCPK PDDLPFSTVVPLKTFYEPGEEITYSC KPGYVSRGGM\RKFICPLTGLWPIN TLKCTPRVCPFAGNLRKMGAVRLIT DFLNYSPTRFSFSLLTWGFILEWAL DS\AKCIEGG
1762	7259	A	1898	19	1215	CQCDSSTMIFSRCSSLFSSFLCHVAI AGRTCPKPDDLPFSTVVPLKTFYEP G\EEITYSCKPGYVSRGGIEESLSCPL \TGTVGPFNTSGNVTPRVCPF\AGIFR KMGGRTLITTF*NYPNTDPVFSLLTL GF*FWNGALDFWPSCTGGKGKW\S P\ELPGLVAPII\CPP\PSIP/TGFATLH VLLRPFRLGNNSPPIGDTAVFECLA HNMAMFG\NDTIT\CTTHGKLDLNY PECRGSKMPPFPHQDPDNGIW*TYP CQNPNTLFTRVKAPHLGLPHDGIFS GMGPRKE\EC*PQTWGKPGSWPLA PSW*KPSLVKGTPVKKRPTVV/YPQ GERVKDSREKFKEWECLHG**KFLS
						FCKNKEKKCSYTEDAQCIDGTIEVP KCFK\EHSSLAFWKT\DAS\DVKPC
1762	7260	A	1899	58	446	RCFREIISSEAT WRT WAS BY THE
1763	7260 7261	$\frac{1}{A}$	1900	1	954	MGEVSGTSDCTDDQCRQVKKALEG
						GKAARGHRSKIKIRFFRPGGLGPGP AITAVAGMPRVYIGRLSYQAREHA V\ERLLNGHAKILEVDLKNGYGFVE FDDLRDADDAVYELNGKDLCGERV IVEHARGPRRDGSYGSGRSGYGYR RSGRDKYGPPTRTEDRLIVEN\LTSR CSWQDLKDYMRQAGEVTYADAHK GRQKMKGVIEFVSYSDMKRALEKL DGTEVNGRKIRLVEDKPGSRRRRSY SRSR\SHSRSRSRSRSRSKSRSRSQSS KSSHSKSRSRSRSRSGSSSKSRSRSSSSSSSSSSSSSSSS
1765	7262	A		3	180	
1766	7263	A	1902	227	440	GMHNVCYVAVNE*FCGFIIR*SLAE RRQIS*EFQLFKFTLCLELILARRAC RESMASPVAGSWSHFPEREF
1767	7264	A	1903	2	438	HEELDTSERKIEFDSASGTYTLYLN GDAHFEEPQSLWNVADLVHQSPPE EKAPLDLSCPQNLFTPK\QEIQWIRI GA\NVS\NFTFAP\STIIFH\LGHA\AM LGLMYVYWTQLNMF\QTLKYLAIL GSVTFLAGNRMLAQQAVKRTAH

SEQ ID NO: of nucleo-tide sequence	SEQ ID NO: of peptide sequence		SEQ 1D NO: in USSN 09/770,160	Nucleotide location of first codon for peptide sequence	location of last codon for last	Amino acid sequence (X=Unknown; *=Stop codon; /=possible nucleotide deletion; \=possible nucleotide insertion)
1768	7265	A	1904		1660	
1768	7265 7266	A	1904	156	1660 2369	PVLKTHPGPQSLPRVPGVPCGGLLE PLSRAEVSPRFGLRRDLLGGMAPPG SSTVFLLALTIIASTWALTPTHYLTK HDVERLKASLDRPFTNLESAFYSIV GLSSLGAQVPDAKKACTYIRSNLDP SNVDSLFYA\AQA\SQGLSGCEISISN ETKDLLLA\AVSE\DSSVYPRSYHAS WQL*SGLLGLSLWAVPKESTQVAL NWLVFKQGKETVL\ATVQALQTAS HLSQQADLRSIVEEIEDLVARLDEL GGLYLQ\FEEGLETTAL\FVAATYKA /LMDH\VGTE\PSIKE\DQVIQLMNAI F\SKKNFES\LSEAFSV\ASG\AAVLS HNRYHVPVVVVPEGSASDTHEQAI LRLQVTNVLSQPLTQATVKLEHAK SVASRATVLQKTSFTP\VGIVFELNF MNVKFSGG*CDF\LVEVEGDNRYIS\ NTVELRVQDPPTEVGITNVDLSTV\ DKDQSIAP\QTTRVTYPAKAKGTFH SAGQATRNFGLVLSSW*DVNTG\AE LTPHQTFVRLHNQKTGPGSGCLFAE PGQQGTCYKFELDTSERKGLNLTSR SGTYTLYLIIG*CQL*RTQILWKCGL MWVIKFP*GKEASFDCLCSQEPFSL
						PKQGNFRHLFPGRP*GRRAPPPWCP NTFTAPESFFGPLL/LCFLRLLWIRD WVPKCLPTFTFCFLSTIIFHPWDML AYAGTSMYVY*TQAQPCSQTLEVP WPILGQCDRFLAGQSGMLAPARQV KRIAAEQSSRLAKYRTLRTAH
1770	7267	A	1906	37	404	PQLSRCRSECMYVNPTVVMTSMGQ ATWSDPHKAKTMLNRIPLGKFAGE SGGSPASVVPAVPVCALGRGGRER WAAASFLYAPDPRPAH\EVEHVVN AILFLLSDRSGMTTGSTLPVEGGFW AC
1771	7268	A	1907	271	1086	YTQCPGIEPVCVDLGDWEATERAL GSVGPVDLLVNNAAVALLQPFLEV TKEAFDR*ACEGGGTSGRGCPGGRS SPNL*PGSVPRPLDPLRVNLRAVIQV SQIVA\RGLI\ARGVPTGPS*NVSSQC FPAGQ*TNHSVLLLPTKGVPLDMLD QG*WAL\ELGPHKLSRCRSGVNA\V NPHSGG*RSMGPGPPWSDPHK\AKI MLNRIP\LGKFAGESEVEHVVNA\IL FLLSDRSGMTTGS\TLPVEGGFWAW LSSLHTPQAPWACFILTPNPSNKT
1772	7269	Α	1908	2	305	ARESGSLVAPRSRPPWEHGLPGEHS *DAPRPHKSPTLPWLPHLHLSKEAL DTHQRSQHE\ECMPLYKFTPTSEKR PQLMLPLPEQQCEQLCRFGSTPVTW A
1773	7270	A	1909	2	529	GTVAACGACYWLLGLMAVRASFE NNCEIGCFAKLTNTYCLVAIGGSEN FYSVFEGELSDTIPVVHASIAGCRIIG RMCVG\TEEILADVLKVEVFRQTVA DQVLVGSYCVFSNQGGLVHPKTSIE

WHAT IS CLAIMED IS:

1. An isolated polynucleotide comprising a nucleotide sequence selected from the group consisting of SEQ ID NO: 1-5497, a mature protein coding portion of SEQ ID NO: 1-5497, an active domain of SEQ ID NO: 1-5497, and complementary sequences thereof.

5

30

- 2. An isolated polynucleotide encoding a polypeptide with biological activity, wherein said polynucleotide hybridizes to the polynucleotide of claim 1 under stringent hybridization conditions.
- 3. An isolated polynucleotide encoding a polypeptide with biological activity, wherein said polynucleotide has greater than about 90% sequence identity with the polynucleotide of claim 1.
 - 4. The polynucleotide of claim 1 wherein said polynucleotide is DNA.
- 15 5. An isolated polynucleotide of claim 1 wherein said polynucleotide comprises the complementary sequences.
 - 6. A vector comprising the polynucleotide of claim 1.
- 20 7. An expression vector comprising the polynucleotide of claim 1.
 - 8. A host cell genetically engineered to comprise the polynucleotide of claim 1.
- A host cell genetically engineered to comprise the polynucleotide of claim 1 operatively
 associated with a regulatory sequence that modulates expression of the polynucleotide in the host cell.
 - 10. An isolated polypeptide, wherein the polypeptide is selected from the group consisting of:
 - (a) a polypeptide encoded by any one of the polynucleotides of claim 1; and
 - (b) a polypeptide encoded by a polynucleotide hybridizing under stringent conditions with any one of SEQ ID NO: 1-5497.
 - 11. A composition comprising the polypeptide of claim 10 and a carrier.
- 35 12. An antibody directed against the polypeptide of claim 10.

5

25

- 13. A method for detecting the polynucleotide of claim 1 in a sample, comprising:
- a) contacting the sample with a compound that binds to and forms a complex with the polynucleotide of claim 1 for a period sufficient to form the complex; and
- b) detecting the complex, so that if a complex is detected, the polynucleotide of claim 1 is detected.
 - 14. A method for detecting the polynucleotide of claim 1 in a sample, comprising:
- a) contacting the sample under stringent hybridization conditions with nucleic acid primers that anneal to the polynucleotide of claim 1 under such conditions;
 - b) amplifying a product comprising at least a portion of the polynucleotide of claim 1; and
- c) detecting said product and thereby the polynucleotide of claim 1 in the sample.
 - 15. The method of claim 14, wherein the polynucleotide is an RNA molecule and the method further comprises reverse transcribing an annealed RNA molecule into a cDNA polynucleotide.
- 20 16. A method for detecting the polypeptide of claim 10 in a sample, comprising:
 - a) contacting the sample with a compound that binds to and forms a complex with the polypeptide under conditions and for a period sufficient to form the complex; and
 - b) detecting formation of the complex, so that if a complex formation is detected, the polypeptide of claim 10 is detected.
 - 17. A method for identifying a compound that binds to the polypeptide of claim 10, comprising:
 - a) contacting the compound with the polypeptide of claim 10 under conditions sufficient to form a polypeptide/compound complex; and
- 30 b) detecting the complex, so that if the polypeptide/compound complex is detected, a compound that binds to the polypeptide of claim 10 is identified.
 - 18. A method for identifying a compound that binds to the polypeptide of claim 10, comprising:

 a) contacting the compound with the polypeptide of claim 10, in a cell, under conditions sufficient to form a polypeptide/compound complex, wherein the complex drives expression of a reporter gene sequence in the cell; and

- b) detecting the complex by detecting reporter gene sequence expression, so that if the polypeptide/compound complex is detected, a compound that binds to the polypeptide of claim 10 is identified.
 - 19. A method of producing the polypeptide of claim 10, comprising,
- a) culturing a host cell comprising a polynucleotide sequence selected from the group consisting of a polynucleotide sequence of SEQ ID NO: 1-5497, a mature protein coding portion of SEQ ID NO: 1-5497, an active domain of SEQ ID NO: 1-5497, complementary sequences thereof and a polynucleotide sequence hybridizing under stringent conditions to SEQ ID NO: 1-5497, under conditions sufficient to express the polypeptide in said cell; and
 - b) isolating the polypeptide from the cell culture or cells of step (a).
 - 20. An isolated polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO: 5498-10994, the mature protein portion thereof, or the active domain thereof.

20

15

- The polypeptide of claim 20 wherein the polypeptide is provided on a polypeptide array.
- 22. A collection of polynucleotides, wherein the collection comprises the sequence information of at least one of SEQ ID NO: 1-5497.
- 25 23. The collection of claim 22, wherein the collection is provided on a nucleic acid array.
 - 24. The collection of claim 23, wherein the array detects full-matches to any one of the polynucleotides in the collection.
- 30 25. The collection of claim 23, wherein the array detects mismatches to any one of the polynucleotides in the collection.
 - 26. The collection of claim 22, wherein the collection is provided in a computer-readable format.

WO 01/79449

PCT/US01/08656

27. A method of treatment comprising administering to a mammalian subject in need thereof a therapeutic amount of a composition comprising a polypeptide of claim 10 or 20 and a pharmaceutically acceptable carrier.

A method of treatment comprising administering to a mammalian subject in need thereof a therapeutic amount of a composition comprising an antibody that specifically binds to a polypeptide of claim 10 or 20 and a pharmaceutically acceptable carrier.